UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/830,497	06/27/2001	Marcin Kuropatwinski	112740-213	8229

29177 7590 05/10/2005

BELL, BOYD & LLOYD, LLC
P. O. BOX 1135
CHICAGO, IL 60690-1135

EXAMINER	
HARPER, V PAUL	

ART UNIT	PAPER NUMBER
2654	

DATE MAILED: 05/10/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/830,497	KUROPATWINSKI, MARCIN	
	Examiner	Art Unit	
	V. Paul Harper	2654	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) 7-12 is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) 7-12 is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☒ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date 10/23/2002.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

DETAILED ACTION

Information Disclosure Statement

1. The Examiner has considered the references listed in the Information Disclosure Statement dated 10/23/2002. A copy of the Information Disclosure Statement is attached to this office action.

Preliminary Amendment

2. The examiner acknowledges the fact the preliminary amendment (submitted on 06/27/2001) is used in the following rejection.

Claim Objections

3. Claim 1 is objected to because of the following informalities:

On line 50 the phrase "to be analysis-by-synthesis coder" should read –to the analysis-by-synthesis coder—

Appropriate correction is required.

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

4. Claim 7 is rejected under 35 U.S.C. 103(a) as being unpatentable over Mogaki et al. ("Text-indicated speaker verification method using PSI-CELP parameters" Security and Watermarking of Multimedia Contents, San Jose, CA, USA, 25-27 Jan 1999), hereinafter referred to as Mogaki, in view of Barnwell et al. ("Speech Coding: A Computer Laboratory Textbook," John Wiley & Sons, Inc, 1996, pp. 127-139), hereinafter referred to as Barnwell, and Sundberg et al. (European Patent Application Publication EP 0817170), hereinafter referred to as Sundberg.

Regarding **claim 7**, Mogaki teaches a method for text-indicated speaker verification using PSI-CELP parameters. Mogaki's method includes the following steps:

- segmenting, in a preparation phase, into first speech signal frames of a given length, a plurality of one of text-dependent and text-independent reference spoken expressions, from a plurality of speakers, which form a speaker-related training statement (Fig. 2, "Input Speech"; Fig. 5, "Speech for Enrollment," §3, system indicated the text which a user should speak; §4.1, each speaker's features are extracted) ;
- supplying the first speech signal frames, in the preparation phase, to an analysis-by-synthesis coder based on linear predictions (Fig. 2, LPC analysis, LPC synthesis filter);
- calculating, in the preparation phase, at least one of a frequency of a respective occurrence of the first parameters in the speaker-related training statement and probability densities with which the first parameters are contained in the speaker-related training statement, the calculation being performed in the analysis-by-synthesis coder

for each of the plurality of speakers and for each first speech signal frame in each case (Fig. 6, "Enrollment Process" PSI-CELP to LSP's to calculation of Variance);

- storing, in the preparation phase, at least one of the calculating frequencies and the probability densities on a speaker-related basis as speaker data (Fig. 6, "Individual Cookbook" with necessary storage);
- calculating, in the usage phase, second probability hits for every third speech signal frame from the calculated third parameters and the speaker data stored for the given speaker in the preparation phase, the second probability hits indicating a probability with which the third parameters have been spoken by the given speaker (Fig. 6, "Verification Process" PSI-CELP to LSP's to calculation of variance);
- combining, in the usage phase, the second probability hits from all the third speech signal frames (Fig. 6, "Calculation of distance"); and
- checking, in the usage phase, to determine whether the combined second probability scores are greater than a predetermined second threshold which identifies the voice of the given speaker, when the combined second probability hits are greater than the predetermined second threshold, the voice of the given speaker is identified, and when the combined second probability scores are less than or equal to the predetermined second threshold, the voice of the given speaker is not identified (Fig. 6, "Verification Process," dist < Threshold?, accept or reject).
- segmenting into third speech signal frames of a given length, in a usage phase, one of a text-dependent and a text-independent used spoken expression of the given

speaker (Fig. 2, "Input Speech"; Fig. 5, "Speech for Verification," §4.1, each speaker's features are extracted);

- supplying, in the usage phase, the third speech signal frames to be analysis-by-synthesis coder (Fig. 2, LPC analysis, LPC synthesis filter);

But Mogaki fails to specifically teach "calculating, **in the preparation phase**, at least one of a first short-term predictor parameter, a first long-term predictor parameter and a first excitation parameter for the coder in the analysis-by-synthesis coder for each of the plurality of speakers and for each first speech signal frame in each case, wherein the parameters form speaker-related training material: and calculating, **in the usage phase**, at least one of a third short-term predictor parameter, a third long-term predictor parameter and a third excitation parameter for the coder, the calculation being performed in the analysis-by-synthesis coder for the given speaker and for every third speech signal frame in each case." However, the examiner contends that this concept was well known in the art, as taught by Barnwell.

In the same field of endeavor, Barnwell teaches basic techniques for speech coding including analysis-by-synthesis coders which include code-excited linear predictive (CELP) coders (p. 127, ¶1). Barnwell also teaches that CELP coders calculate short-term, long-term, and excitation parameters (§7.8).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Mogaki by specifically providing the specifics of CELP coding, as taught by Barnwell, because it is well known in the art at the time of invention that these are the standard techniques for calculated CELP parameters.

In addition, Mogaki in view of Barnwell teaches much of the material described in the limitations a) through f) below and where applicable are rejected for the same reasons given above, but Mogaki in view of Barnwell does not specifically teach that during the simulation usage phase training results are combined until a particular level of performance is reached. The simulated usage limitations are listed as follows:

a) segmenting, **in a simulated usage phase of the training phase**, into second speech signal frames of a given length, one of a text-dependent and a text independent simulation spoken expression of a given speaker;

b) supplying, **in the simulated usage phase,** the second speech signal frames to the signal-by-synthesis coder;

c) calculating, **in the simulated usage phase**, at least one of a second short-term predictor parameter, a second long-term predictor parameter and a second excitation parameter for the coder, the calculation being performed in the analysis-by-synthesis coder for the given speaker and for every other speech signal frame in each case;

d) calculating, **in the simulated usage phase**, first probability hits for every other speech signal frame from the calculated second parameters and the speaker data stored for the given speaker in the preparation phase, the probability hits indicating a probability with which the second parameters match the first parameters;

e) combining, **in the simulated usage phase**, the first probability scores from all the second speech signal frames;

f) checking, in the simulated usage phase, to determine whether the combined first probability scores are greater than a predetermined first threshold which confirms the voice of the given speaker, when the combined first probability scores are greater than the predetermined first threshold, the voice of the given speaker is confirmed, and when the combined first probability scores are less than or equal to the predetermined first threshold, the preparation phase continues for further reference spoken expressions by the given speaker until the voice of the given speaker is confirmed.

However, the examiner contends that these concepts were well known in the art, as taught by Sundberg.

In the same field of endeavor, Sundberg teaches a method for the adaptation of models used in speaker verification systems. In particular, Sundberg teaches the training [combining] of speaker verification models until the performance [checking the combined scores] reaches a particular level [threshold], in particular f), above, (abstract, col. 2, lines 15-22, col. 4, lines 1-15, the complex models can be trained [during a simulated usage stage, in a) through f), above] until they are ready to be put into use).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Mogaki in view of Barnwell by specifically providing the features, as taught by Sundberg, because it is well known in the art at the time of invention for the purpose of dynamically adapting a model until it reaches the desired level of performance (Sundberg, col. 1, line 54 through col. 2, line 6).

5. Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Mogaki in view of Barnwell and Sundberg, and further in view of Gersho et al. (U.S. Patent 6,233,550), hereinafter referred to as Gersho.

Regarding **claim 8**, Mogaki in view of Barnwell and Sundberg teaches everything claimed, as applied above (see claim 7), but Mogaki does not specifically teach "one of a harmonic vector excited predictive coder and a waveform interpolating coder is used as a parametric coder." However, the examiner contends that this concept was well known in the art, as taught by Gersho.

In the same field of endeavor, Gersho discloses a method for hybrid coding of speech at 4kbps. In addition, Gersho teaches that harmonic coders excel at low bit rates (col. 3, line 55 through col. 4, line 20).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Mogaki in view of Barnwell and Sundberg by specifically providing the features, as taught by Gersho, because it is well known in the art at the time of invention for the purpose of reducing the bit rate (Gersho, col. 4, lines 15-20).

6. Claim 9 is rejected under 35 U.S.C. 103(a) as being unpatentable over Mogaki in view of Barnwell and Sundberg, and further in view of Hagen et al. (U.S. Patent 6,182,030), hereinafter Hagen.

Regarding **claim 9**, Mogaki in view of Barnwell and Sundberg teaches everything claimed, as applied above (see claim 7). But Mogaki does not specifically teach "an LPAS coder is used as the analysis-by-synthesis coder." However, the examiner contends that this concept was well known in the art, as taught by Hagen.

In the same field of endeavor, Hagen discloses a method for enhanced coding to improve coded communication signal. In addition, Hagen teaches the linear-prediction based analysis-by-synthesis (LPAS) paradigm (col. 1, lines 30-37).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Mogaki in view of Barnwell and Sundberg by specifically providing the features, as taught by Hagen, because it is well known in the art at the time of invention for the purpose of improved coding at the rates between 5 and 20 kb/s (col. 1, lines 30-40).

7. Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Mogaki in view of Barnwell and Sundberg, and further in view of Acero et al. (U.S. Patent 5,535,305), hereinafter referred to as Acero.

Regarding **claim 10**, Mogaki in view of Barnwell and Sundberg teaches everything claimed, as applied above (see claim 7). But Mogaki does not specifically teach the step of "quantizing at least one of the frequencies and the probability densities using a vector quantizer having a specific and considerably reduced number of bits."

However, the examiner contends that this concept was well known in the art, as taught by Acero.

In the same field of endeavor, Acero discloses a technique for sub-partitioned vector quantization of probability density functions to reduce the memory requirements (abstract; col. 2, lines 24-28, applied to speech recognition).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Mogaki in view of Barnwell and Sundberg by specifically providing the features, as taught by Acero, because it is well known in the art at the time of invention for the purpose of reducing storage requirements (col. 1, lines 8-11).

8. Claims 11 and 12 are rejected under 35 U.S.C. 103(a) as being unpatentable over Mogaki in view of Barnwell and Sundberg, and further in view of Boll ("Suppression of Acoustic Noise in Speech Using Spectral Subtraction," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. ASSP-27, No. 2, April 1979), hereinafter referred to as Boll.

Regarding **claims 11 and 12**, Mogaki in view of Barnwell and Sundberg teaches everything claimed, as applied above (see claim 7). But Mogaki does not specifically teach the step of "entering noise which is known to the speaker identification system when the spoken expression of a speaker is entered into the speaker identification

system, and subtracting the entered noise internally, before the segmentation, from the recording of the speakers voice."

However, the examiner contends that these concepts were well known in the art, as taught by Boll.

In the same field of endeavor, Boll teaches the suppression of acoustic noise in speech using spectral subtraction applied to speech recognition or speaker authentication systems (abstract). Boll Further teaches that words can be recorded in a noisy [helicopter] environment and the noise can be subtracted before further processing (p. 119, §C).

Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Mogaki in view of Barnwell and Sundberg by specifically providing the features, as taught by Boll, because it is well known in the art at the time of invention for the purpose of reducing noise during pre-processing in speaker authentication systems (Boll, abstract).

Citation of Pertinent Art

9. The following prior art made of record but not relied upon is considered pertinent to the applicant's disclosure:

- Beith et al. (U.S. Patent 6,449,496 B1) discloses a voice recognition user interface for telephone handsets that includes the repetition of the enrollment procedure until the words match.
- Rissanen (U.S. Patent 5,430,827) discloses a voice recognition password verification system.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to V. Paul Harper whose telephone number is (571) 272-7605. The examiner can normally be reached on M-F.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

05/05/2005

V. Paul Harper
Patent Examiner
Art Unit 2654